

Technical Bulletin

Part No. 00D-TB038

DataStage Row Merger

This technical bulletin describes DataStage Row Merger. This stage allows you to combine a number of input columns into a single column and output it.

© 2003 Ascential Software Corporation. All rights reserved. Ascential, Ascential Software, DataStage, MetaStage, and MetaBroker are trademarks of Ascential Software Corporation or its affiliates and may be registered in the United States or other jurisdictions. Adobe Acrobat is a trademark of Adobe Systems, Inc. Microsoft, Windows, Windows NT, and Windows Server are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company, Ltd. Other marks mentioned are the property of the owners of those marks.

This product may contain or utilize third party components subject to the user documentation previously provided by Ascential Software Corporation or contained herein.

Printing History

First Edition for Release 1.0, December 2003

How to Order Technical Documents

To order copies of documents, contact your local Ascential subsidiary or distributor, or call our main office at (508) 366-3888.

Documentation Team: Mandy deBelin and Marie E. Hedin

Introduction

This technical bulletin describes the following for Release 1.0 of the DataStage Row Merger stage for DataStage Release 7.0.1:

- [Functionality](#)
- [Installation](#)
- [Reading data](#)
- [Formatting data](#)
- [Writing data](#)

Row Merger reads data one row at a time from an input link. It merges all the columns into a single string of a specified format. It then writes the string on a given column of the output link. The stage can have a single input link and a single output link.

In normal operation of Row Merger, each input row with multiple columns results in an output row of a single column. The stage also offers concatenation facilities, however. These facilities allow you to concatenate the result of each input row into a single string which is output when the stage detects an end-of-data (EOD) or end-of-transmission (EOT) signal (that signifies no more input rows are expected).

Note: The row merger stage is similar to the server sequential file stage. The difference is that, while the sequential file stage writes to a file, Row Merger outputs to a link.

This bulletin describes the user interface for Row Merger.

Functionality

Row Merger has the following functionality and benefits:

- Reads one row at a time, merges all the columns from a row into a single string of a specified format, and then writes the string to a given column of the output link.
- Supports aggregation of multiple rows of data.
- Supports NLS (National Language Support). The stage writes what it reads without interpretation or conversion.

The following functionality is not supported:

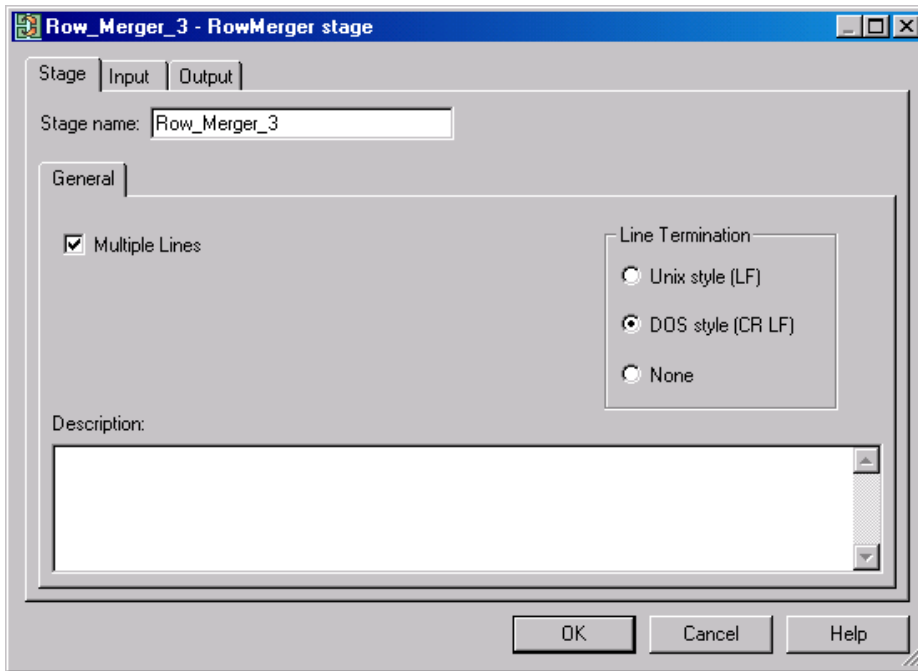
- Support for MetaStage.

Installing the Plug-in

For instructions and information supporting the installation, see *DataStage Plug-In Installation and Configuration Guide*.

Stage Page General Tab

The **General** tab of the **Stage** page gives access to the concatenation facilities of Row Merger.



The **General** tab contains the following fields:

- **Multiple Lines.** This determines whether Row Merger concatenates input rows into a single output row, or whether it outputs each input row as a separate output row. Select **Multiple Lines** to have the rows concatenated. By default it is not selected.
- **Line Termination.** This setting is only available if you have chosen the **Multiple Lines** option to specify that the stage is concatenating input rows. It specifies the character(s) that will be placed as a delimiter between the concatenated rows when they are output in a single row. Choose from the following settings:
 - **Unix Style (LF).** Places a linefeed character as a delimiter between each merged row.
 - **DOS style (CR LF).** Places a carriage return character and a linefeed character as a delimiter between each merged row.
 - **None.** Does not place a delimiter between the merged rows.
- **Description.** Enter an optional description of the stage.

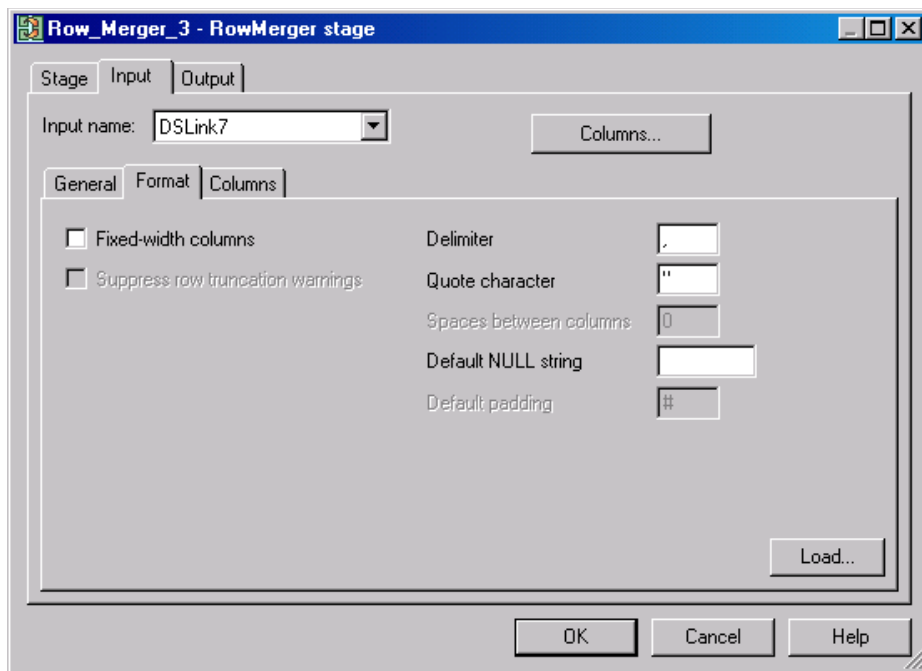
Inputs Page

The **Inputs** page contains various tabs that describe the rows of data being input to Row Merger.

The **General** tab contains a description field that allows you to enter an optional description of the input link. The **Format** tab and **Columns** tab are described below.

Format Tab

This tab allows you to specify how the data read in individual columns in each input row will be formatted before being output in a single column.



The tab contains the following fields:

- **Fixed-width columns.** Select this check box to output the data in fixed-width format. The width of each field is taken from the SQL display size of the input columns (set in the **Display** column in the Columns grid on the **Inputs** page **Columns** tab). This option is cleared by default.
- **Suppress row truncation warnings.** This option is only available when you have selected **Fixed-width columns**. If the input rows contain more columns that you have defined on the **Columns** tab, you will normally receive warnings about overlong rows when the job is run. If you want to suppress these messages (for example, you may only be interested in the first three columns and happy to ignore the rest), select this checkbox.
- **Delimiter.** This option is not available if you have selected **Fixed-width columns**. It specifies the delimiter used to separate the data fields in the output data that have been derived from the input columns. By default this field contains a comma. You can enter a single printable character or a decimal or hexadecimal number to represent the ASCII code for the character you want to use. Valid ASCII codes are in the range 1 to 253. Decimal values 1 through 9 must be preceded with a zero. Hexadecimal values must be prefixed with &h. Enter 000 to suppress the delimiter.
- **Quote Character.** This option is not available if you have selected **Fixed-width columns**. Specifies the character used to enclose strings. By default this field contains a double quotation mark. You can enter a single printable character or a decimal or hexadecimal number to represent the ASCII code for the character you want to use. Valid ASCII codes are in the range 1 to 253. Decimal values 1 through 9 must be preceded with a zero. Hexadecimal values must be prefixed with &h. Enter 000 to suppress the quote character.

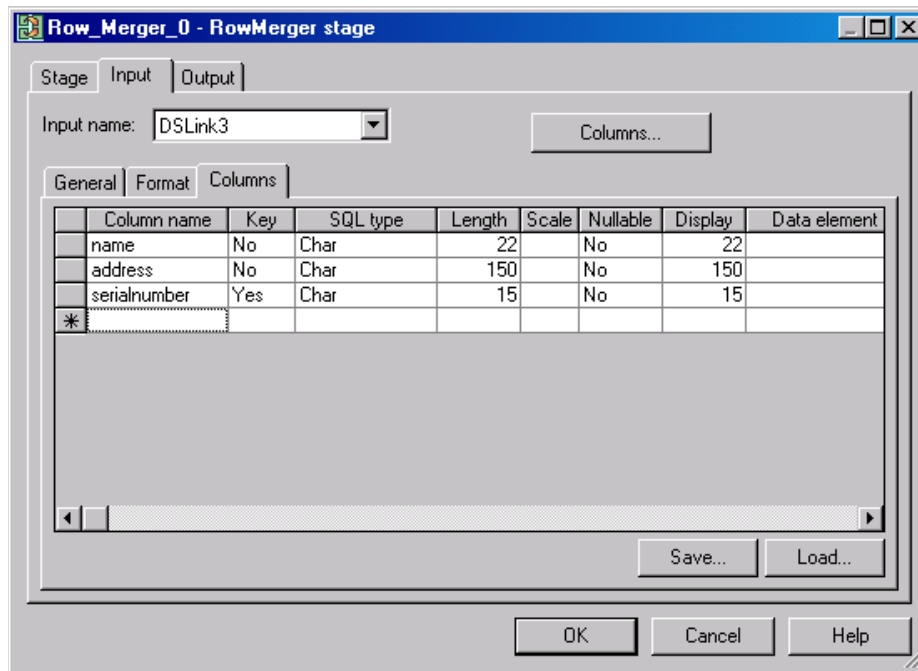
- **Spaces between columns.** This option is only available if you have selected **Fixed-width columns**. Contains a number to represent the number of spaces used between columns. By default this is 0.
- **Default NULL string.** Contains characters which, when encountered in an input row, are interpreted as the SQL null value.
- **Default padding.** This option is only available if you have selected **Fixed-width columns**. Contains the character used to pad missing columns. This is # by default, but can be set to another character here.

The **Format** tab also has a **Load** button. If you have table definitions that include format information, you can load the format details from these table definitions directly onto the **Format** page:

1. Click **Load**. The **Load Table Definitions** dialog box appears.
2. Browse for the table definition containing the format you want to load.
3. Click **OK**. The format details are loaded.

Columns Tab

The entries in the columns grid specify the format of the data being read from the input rows.

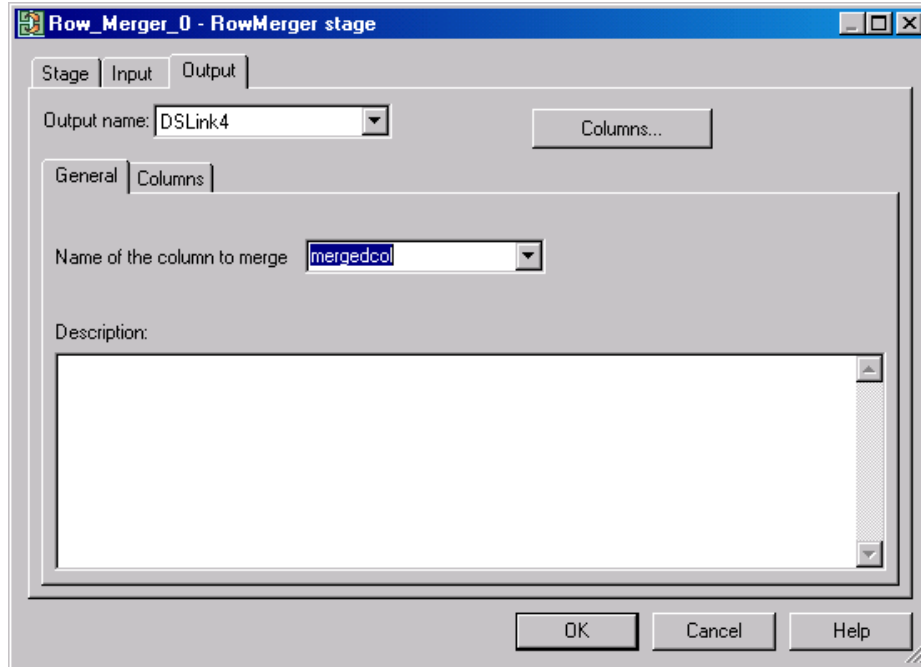


The grid has the standard fields that all column definitions have.

Outputs Page

The **Outputs** page contains various tabs that describe the data being output by Row Merger.

General Tab

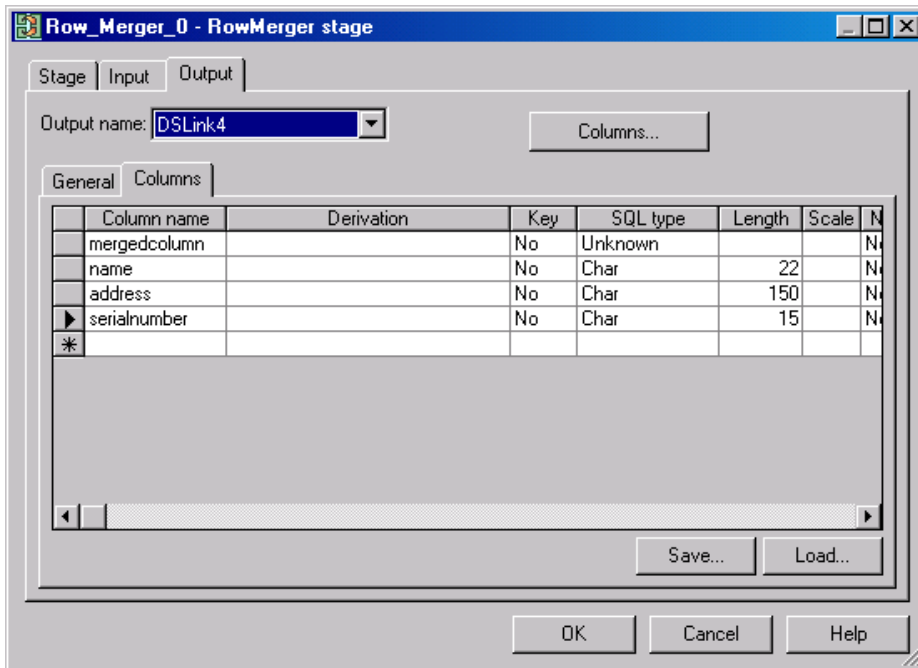


The **General** tab contains the following fields:

- **Name of the column to merge.** A drop-down list contains a list of the defined output columns for this stage. Choose the column that you want to output the merged data in.
- **Description.** Allows you to enter an optional description of the output link.

Columns Tab

The entries in the columns grid specify the format of the data being written to the output link.



The grid has the standard fields that all column definitions have. The Derivation field is not used.

At the least you must define a column to carry the merged data. You can also define additional columns to carry the data as input to the stage. In the example above mergedcolumn carries the result of the merge operation, i.e., a single column containing all the columns in the input data. The remaining columns carry the data from the original name, address, and serialnumber columns. For this feature to work, the output columns must have the same names as the input columns.